

EXHIBIT G

Alex Frawley

From: Alex Frawley
Sent: Thursday, October 24, 2024 10:11 AM
To: Levy, Max I.; KVP-OAI; openaicopyrightlitigation.lwteam@lw.com; OpenAICopyright
Cc: #NewYorkTimes-Microsoft-FDBR; NewYorkTimes_Microsoft_OHS@orrick.com; NYT-AI-SG-Service@simplelists.susmangodfrey.com; dailynews-ai-rfem@rothwellfigg.com
Subject: RE: NYT / OAI's Responses to RFPs

Hi Max,

Thanks for these responses. Let's meet and confer today at 3:30 ET. We would like to discuss RFPs 6, 11, and 18.

Best,
Alex

From: Levy, Max I. (via NYT-AI-SG-Service list) <NYT-AI-SG-Service@simplelists.susmangodfrey.com>
Sent: Wednesday, October 23, 2024 7:47 PM
To: Alex Frawley <AFrawley@susmangodfrey.com>; KVP-OAI <KVPOAI@keker.com>; openaicopyrightlitigation.lwteam@lw.com; OpenAICopyright <OpenAICopyright@mofo.com>
Cc: #NewYorkTimes-Microsoft-FDBR <NewYorkTimes-Microsoft-FDBR@faegredrinker.com>; NewYorkTimes_Microsoft_OHS@orrick.com; NYT-AI-SG-Service@simplelists.susmangodfrey.com; dailynews-ai-rfem@rothwellfigg.com
Subject: RE: NYT / OAI's Responses to RFPs

EXTERNAL Email

Hi Alex, please see our response inline below in [blue](#). We're available to meet and confer on these issues tomorrow (Thursday) after 11am PT and Monday between 1pm and 5pm PT.

Best,
Max

MAX LEVY

Associate | Morrison & Foerster LLP
755 Page Mill Road | Palo Alto, CA 94304-1018
P: +1 (650) 813-4063

From: Alex Frawley <AFrawley@susmangodfrey.com>
Sent: Tuesday, October 22, 2024 8:01 PM
To: Levy, Max I. <MLevy@mofo.com>; KVP-OAI <KVPOAI@keker.com>; openaicopyrightlitigation.lwteam@lw.com; OpenAICopyright <OpenAICopyright@mofo.com>
Cc: #NewYorkTimes-Microsoft-FDBR <NewYorkTimes-Microsoft-FDBR@faegredrinker.com>; NewYorkTimes_Microsoft_OHS@orrick.com; NYT-AI-SG-Service@simplelists.susmangodfrey.com; dailynews-ai-rfem@rothwellfigg.com
Subject: RE: NYT / OAI's Responses to RFPs

External Email

Counsel,

Thank you for the meet and confer on Friday, October 17 to discuss various disputes. We write to follow up on the call and the parties' disputes, as reflected in the conference agenda.

For many of the disputed RFPs (Issue 6 below), we request that OpenAI confirm by close of business tomorrow whether it will, or will not, agree to produce the requested documents to ensure that we have adequate time to tee up any remaining disputes in advance of the upcoming conference. We flag below the specific RFPs for which this request applies.

1. Dkt. 128 (dispute regarding GPT, GPT-2, and GPT-3)

OpenAI's responses to The Times's Interrogatories defined the at-issue models as limited to "only the models used for ChatGPT," and OpenAI incorporated that objection into each of its Interrogatory responses. As we explained on the call, given the parties' agreement that GPT, GPT-2, and GPT-3 will be subject to discovery, and the fact that OpenAI will be producing documents about these models, OpenAI should revisit its Interrogatory responses and supplement them where appropriate to address these earlier models. For some Interrogatories, the substance of OpenAI's answer may not change; for example, for Interrogatory No. 1, there may not be additional individuals "with knowledge of Defendants' Training Datasets" for these earlier models apart from the individuals whom OpenAI has already identified for "the models used for ChatGPT." But for every Interrogatory, OpenAI should supplement its response to either (A) provide additional information for these earlier models, if appropriate, or (B) confirm that the Interrogatory response remains complete and accurate, including with respect to GPT, GPT-2, and GPT-3. Put differently, OpenAI should no longer limit its responses to "only the models used for ChatGPT."

Please confirm that OpenAI will (1) amend its interrogatory responses to include any responsive information about GPT, GPT-2, and GPT-3, and (2) refrain from objecting to future discovery requests on the ground that GPT, GPT-2, and GPT-3 are not subject to discovery.

In light of OpenAI's agreement that GPT, GPT-2 and GPT-3 will be subject to discovery, OpenAI will, without conceding the relevance of these models, agree to review its interrogatory responses and supplement responses where appropriate. Consistent with its obligations under the federal rules, OpenAI will continue to ensure that its responses are complete and accurate to the best of OpenAI's understanding. If The Times believes it appropriate to supplement particular interrogatory responses (or not), The Times is free to share that information.

2. Dkt. 141 (dispute regarding The Times's RFP 6 and 11)

Please let us know when OpenAI is available to meet and confer about these disputes, which were briefed in the agenda. The News Plaintiffs' positions are memorialized in the agenda.

[See email above for proposed times to meet and confer regarding these disputes.](#)

3. Dkt. 204 (dispute regarding additional OpenAI custodians)

Please let us know if OpenAI is still taking the position that none of the eight at-issue employees should be custodians. Please also let us know whether OpenAI has produced any custodial documents since the September 12 conference, or if it intends to do so before the October 30 conference.

[OpenAI's position as to the eight at-issue custodians is memorialized in the agenda.](#)

[OpenAI intends to produce additional custodial documents before the October 30 conference.](#)

4. Dkt. 228 (dispute regarding OpenAI's search terms)

The News Plaintiffs in the conference agenda requested that the Court set a target of approximately 550,000 to 600,000 additional documents for OpenAI to review, with the parties to then work out the exact terms between themselves. Does OpenAI agree to this proposal? If not, does OpenAI have a counter-proposal for an approximate target number of documents?

OpenAI remains open to finding a compromise position with Plaintiffs. In our experience, 10,000 documents per witness is the upper limit of what we've seen in cases our team has collectively litigated. Accordingly, we are willing to make a counter-proposal that yields approximately 240,000 total documents. If The Times has support for a higher number, please share it so that we may consider your position.

5. Dispute regarding the scope of OpenAI's products subject to discovery

We discussed this dispute during our call on October 14, which was led by Mark Rawls. I followed up on October 17 by email to make a proposal on behalf of the News Plaintiffs, which I paste below. Please respond with OpenAI's position on this proposal.

"The News Plaintiffs' proposal is that the scope of the at-issue products for discovery be: ChatGPT, Browse, the OpenAI API, Copilot, SearchGPT, and any current or anticipated GPT-based product (including CustomGPTs) that relate to news content (including using retrieval augmented generation to retrieve news content), product recommendations, or recipe recommendations."

As The Times wrote in the agenda, "OpenAI's main concern was having a fixed set of products to work with." This continues to be our main concern and The Times's proposal to include "anticipated GPT-based product[s]" does not alleviate that concern. How does The Times define an "anticipated" product? We continue to take the position from our agenda: OpenAI is "willing to consider broadening the definition—even to products that Plaintiffs have never identified in any of their complaints. The issue still being negotiated involves OpenAI's request that the list of products and services be discrete and reasonable in number." Is The Times willing to remove "anticipated" products from the scope of the products-at-issue?

6. RFP Disputes

During the October 17 call, we discussed OpenAI's responses to New York Times RFPs 18, 26-27, 44, 47-48, 51-53, 74-81, and 84-85, all of which were briefed in the agenda. Below, I summarize next steps for these RFPs. As noted above, for some of these RFPs, we ask that OpenAI by close of business tomorrow confirm that they will produce the requested documents. The Times will otherwise move to compel. This request applies to RFPs 18, 26, 27, 44, 47, 51, and 74-81, and 84.

RFP 18

OpenAI explained that its objection to the News Plaintiffs' proposal is that the wording could be interpreted to mean that OpenAI is admitting to removing CMI. To account for that concern, how about this revision to the News Plaintiffs' proposal: "documents concerning **any** removal of copyright management information before, during, or after training of Defendants' Generative AI Model(s), including: information regarding authorship, date of publication, publishing entity, title, copyright notices, and terms and conditions for use of works"? The News Plaintiffs can also clarify that the portion of the request referring to "after training" refers to any post-training processes, as well as any removal of CMI in connection with outputs provided to users, including for retrieval augmented generation ("RAG").

Please confirm by close of business tomorrow that OpenAI will agree to produce the requested documents.

OpenAI will, subject to a reasonable search, agree to produce non-privileged documents in our possession, custody, or control concerning the removal, if any, of copyright management information from datasets on which the relevant models were trained, including: information regarding authorship, date of publication, publishing entity, title, copyright notices, and terms and conditions for use. This response provides The Times with precisely the information it requested in its October 9 email. (“This RFP seeks documents concerning the removal of any content from Training Datasets, including information regarding authorship, date of publication, publishing entity, title, copyright notices, and terms and conditions of works.”)

Separately, we’re surprised to see the limitation as to Training Datasets be removed—and to learn for the first time that the scope of requested information now includes RAG. In fact, RFP 18 as-served specifically requested “the removal of any content . . . from Training Datasets.” As The Times knows, the use of RAG takes place after the training (pre-training and post-training) is completed, so it cannot “remov[e] . . . any content . . . from Training Datasets.” We are open to meet and confer about the production of this newly requested information—particularly if there is an RFP requesting it. But as a factual matter, any alleged removal of CMI via RAG is not within the scope of RFP 18 as written.

Our proposed response to RFP 18 provides The Times with what it asked for on October 9. Does this resolve the dispute?

RFPs 26-27

As discussed on the call, for RFP 26, OpenAI’s proposal is deficient because it omits the use of web crawlers for RAG. In addition, OpenAI told us that its response would cover OpenAI’s use (if any) of the specific tools listed in the RFP, including GPTBOT, ChatGPT-User, OAI-SearchBot, CommonCrawl bot, and crawlers used by Microsoft or developers of CustomGPTs. Please confirm that is the case by close of business tomorrow. Please also confirm that OpenAI will produce documents regarding the use of web crawlers for RAG.

We continue to believe that our prior proposal is sufficient for the reasons discussed on the call. Further, we note that OpenAI’s response to RFP No. 12 (“how OpenAI uses RAG in connection with ChatGPT”) already includes documents sought by the Times’s latest proposal. However, to try to resolve this dispute, OpenAI agrees to conduct a reasonable search for and produce non-privileged documents in its possession, custody or control discussing OpenAI’s use of web crawlers to copy articles published by The Times in connection with training the relevant models and ChatGPT. As defined, “ChatGPT” includes Browse, which is ChatGPT’s RAG functionality, so this proposal addresses the Times’s concern about the omission of web crawlers for RAG.

That said, we cannot agree to the Times’s addition of “crawlers used by Microsoft or developers of CustomGPTs” to this RFP (which, as served, did not mention either sets of products). In the first place, these terms are wildly overbroad and would encompass an endless number of third-party products that have no apparent relevance to this case. For example, the term “crawlers used by Microsoft” could include the manner in which Bing accesses Times content to produce its own search results. Moreover, because these crawlers are operated by third parties, our understanding is that OpenAI has no control over the manner in which they function. If The Times has information to suggest otherwise, please share it so that we may consider it.

For RFP 27, you stated on the call that you were under the impression OpenAI would be withdrawing its “sufficient to show” limitation for this RFP. Please confirm that is correct, and that OpenAI will be producing custodial documents in response to this request. In addition, the News Plaintiffs cannot agree to OpenAI’s proposal to limit productions to OpenAI’s “approach to” robots.txt, which could be interpreted to narrow the request. OpenAI should produce documents concerning its reliance on, use of, or understanding of robot.txt. Please confirm by close of business tomorrow that OpenAI will do so.

As we explained on the meet-and-confer, OpenAI interprets the term “approach to robots.txt” to be broader than and inclusive of The Times’s proposed language of “reliance on, use of and understanding of robots.txt.” We therefore do not understand what this dispute is about. Moreover, we’re surprised to learn that the Times now takes issue with the use of the term “approach to,” since The Times suggested during the October 18 meet and confer that it did not object to our interpretation of that phrase as broader than The Times’s proposed language. In any event, we reiterate here that to the extent relevant, non-privileged documents concerning OpenAI’s “reliance on, use of, or understanding of robot.txt” exist and can be located pursuant to a reasonable search, they will be included within OpenAI’s agreement to produce responsive, non-privileged documents in its possession, custody, and control, if any, concerning OpenAI’s approach to robots.txt, that it locates pursuant to a reasonable and diligent search, and that custodial documents will be included in that search.

RFPs 44, 53

For RFP 44, we explained our concern with OpenAI’s proposal to produce documents regarding its “approach to paywalls with respect to the training data for the relevant models and products.” For example, that wording is too narrow because it could be limited to OpenAI’s policies and procedures. This proposal also appears to omit RAG. Please confirm that OpenAI will instead agree to produce documents in response to the RFP as drafted: “Documents concerning any attempts to use Defendants’ Generative AI Products and Services to circumvent paywalls or other access control measures.” As explained on the call, the News Plaintiffs can also clarify that this request seeks documents concerning attempts by users to use Defendants’ Generative AI Products and Services to circumvent paywalls or other access control measures.

Please confirm by close of business tomorrow that OpenAI will produce the requested documents.

Similar to RFP 27, OpenAI interprets the term “approach to paywalls” to be broader than and inclusive of The Times’s proposed language. We reiterate here that to the extent relevant, non-privileged “[d]ocuments concerning any attempts to use [the relevant products] to circumvent paywalls or other access control measures” exist and can be located pursuant to a reasonable search, they will be included within OpenAI’s agreement to produce responsive, non-privileged documents in its possession, custody or control, if any, discussing OpenAI’s approach to paywalls with respect to the relevant models and products.

We also confirm that OpenAI’s current offer includes RAG. The scope is “with respect to the relevant models and products.” The parties have already clarified – as reflected in OpenAI’s October 11 supplemental interrogatory response – that ChatGPT (a relevant product) includes Browse; and to the extent RAG is implicated by any other “relevant models and products” it would be within the scope of the proposed response. Moreover, to the extent any documents related to RAG’s interaction with paywalls exist, OpenAI has already agreed to produce them in its response to RFP 12 (“how OpenAI uses RAG in connection with ChatGPT”). Put differently, we reiterate that RAG is not omitted from OpenAI’s offer.

For RFP 53, we can accept your offer, as memorialized in the agenda on page 31, and copied below:

“OpenAI has already agreed to conduct a reasonable search across its existing custodians and produce non-privileged documents in its possession, custody, or control, related to Browse with Bing’s interactions with paywalls. OpenAI has also agreed to produce documents, to the extent they exist, regarding how paywalls may have related to the decision to suspend Browse with Bing and/or Bing Chat, if at all.”

RFPs 45, 47-48, 51-52

For RFP 45, OpenAI represented that it is still investigating its response. Please provide the results of that investigation by October 25.

We understand that the parties are engaged in an ongoing discussion on OpenAI's collection and production of log files for its GPT products and services. *See* Emails between Paven Malhotra and Jen Maisel "RE: NYT/NYDN - Requests for Inspection to OpenAI." For an update on the investigation, we direct The Times to that discussion.

For RFP 47, we explained how OpenAI's proposal is deficient because the use of "analyzing or summarizing" could be construed to make the response more narrow than the request, which seeks documents "concerning" user engagement metrics. The response is also deficient because it appears to be limited to click-through rates while omitting the other types of metrics listed in the RFP, including session duration, prompts or queries per session, and bounce rates.

Please confirm by close of business tomorrow that OpenAI will agree to produce these documents.

The News Plaintiffs can also clarify that "fallback rates" refer to search methods that present related results when there are limited results for a specific query. For example, instead of displaying only a small number of products in response to a specific search, a platform can display related products the user might also link.

OpenAI agrees to produce responsive, non-privileged documents in its possession, custody, and control, if any, concerning the following user engagement metrics: time-spent, average number of prompts, click-through data, and bounce rate.

We appreciate you following up on the meaning of "fallback rates." But if "fallback rates" are "search methods," we do not understand how they could be relevant to the request. First, "search methods" are not a category of "user engagement metrics, events, and analytics data collected during each user session" sought by the RFP. The Times has repeatedly (dating back to August) emphasized that this RFP is seeking user engagement metrics. Is it now expanding the scope to include a "search method"? If so, we'd request that The Times elaborate on its relevance at meet and confer. But given that our offer now encompasses every user engagement metric identified by The Times in RFP 47, we're hopeful that our response resolves any dispute as to this RFP.

For RFP 48, please confirm that OpenAI does not in the ordinary course of business maintain information about the number of daily active users separated into new users and repeat users. As discussed on the call, the News Plaintiffs can clarify that we are not asking OpenAI to produce a list of subscriber's names. We are instead seeking aggregate statistics regarding dates of subscription and changes in subscriber numbers.

As discussed on the meet and confer, OpenAI agrees to produce non-privileged documents showing the number of subscribers and daily active users of each of the relevant products, from 2021 to the present, in the manner and at the frequency such information is kept in the ordinary course of business.

We understand the response's reference to "subscribers and daily active users" to reflect what you mean by "new users and repeat users." If that's not the case, please let us know so we can better understand your position. But based on the last meet and confer, we believe that OpenAI's updated response resolves the dispute.

For RFP 51, OpenAI's proposed limitation to "analyzing or summarizing" click-through rates presents the same problems addressed above for RFP 47. OpenAI should instead commit to produce documents "concerning" the issues listed in the RFP. The News Plaintiffs can clarify that while RFPs 51 and 47 are similar, they are different because while RFP 47 focuses on user engagement metrics, RFP 51 focuses on the "behavior of users," which is broader.

Please confirm by close of business tomorrow that OpenAI will agree to produce the requested documents.

OpenAI agrees to produce non-privileged documents in its possession, custody, or control concerning OpenAI user click-through data for the relevant products, subject to a reasonable search. Additionally, OpenAI agrees to

produce non-privileged documents in its possession, custody, or control discussing OpenAI's efforts, if any, to modify clickthrough rates for the relevant products, subject to a reasonable search. Given that we have adopted Plaintiffs' use of the term "concerning," we believe this resolves the dispute.


As to the difference between "user engagement metrics" and the "behavior of users," if this is an issue preventing resolution of the dispute, we're eager to meet and confer. So far as we understand, the only "user behavior" OpenAI can know about is what it is able to measure – and those measurements are user engagement metrics. (User Pilot: User Engagement ("Measuring user engagement involves tracking various metrics that provide insights into user behavior.")) If The Times believes our response is deficient because it does not sufficiently address "user behavior," we are open to further amending our response, but need to better understand The Times's position.

For RFP 52, OpenAI represented that it is still investigating its response. Please provide the results of that investigation by October 25.

We understand that the parties are engaged in an ongoing discussion on OpenAI's collection and production of log files for its GPT products and services. *See* Emails between Paven Malhotra and Jen Maisel "RE: NYT/NYDN - Requests for Inspection to OpenAI." For an update on the investigation, we direct The Times to that discussion.

RFPs 74-81

RFPs 74-80 each seek documents



Please confirm by close of business tomorrow that OpenAI will agree to produce the requested documents.



Please also confirm by close of business tomorrow whether OpenAI has conducted a specific investigation into the information sought by RFP 81, which likewise cited specific documents produced in discovery, and that OpenAI will produce documents in response to this request as well.

We can confirm that the information sought by RFP 81 is within the scope of our proposed response and that we will produce responsive documents based on a reasonable search.

RFPs 84-85

For RFP 84, OpenAI wrote in the agenda that it is using “DNYuz” as a search term. We asked on the call for more specifics about the term and hit counts. Can you please provide that information? Was this term used for every custodian? What were the hit counts?

Given that OpenAI has apparently run at least some form of “DNYuz” as a search term, please confirm by close of business tomorrow that OpenAI will produce documents it finds in response to this RFP. Depending on your responses to the above questions, the parties can work out any issues relating to the exact search terms used for this RFP.

We understand that The Times is seeking hit reports associated with the “DNYuz” search term. Our understanding is that The Times has not included “DNYuz” as a search term in its proposals yet, but we anticipate that you’ll do so in a counterproposal. Should The Times do so, OpenAI will of course provide the hit report pursuant to the ESI protocol.

Finally, we will write separately regarding RFP 85.

Best,
Alex

From: Alex Frawley (via NYT-AI-SG-Service list) <NYT-AI-SG-Service@simplelists.susmangodfrey.com>
Sent: Thursday, October 17, 2024 9:14 AM
To: Levy, Max I. <MLevy@mofo.com>; KVP-OAI <KVPOAI@keker.com>; openaicopyrightlitigation.lwteam@lw.com; OpenAICopyright <OpenAICopyright@mofo.com>
Cc: #NewYorkTimes-Microsoft-FDBR <NewYorkTimes-Microsoft-FDBR@faegredrinker.com>; NewYorkTimes_Microsoft_OHS@orrick.com; NYT-AI-SG-Service@simplelists.susmangodfrey.com; dailynews-ai-rfem@rothwellfigg.com
Subject: RE: NYT / OAI's Responses to RFPs

EXTERNAL Email

We can do 4 ET. I will send an invite. Thanks!

From: Levy, Max I. <MLevy@mofo.com>
Sent: Wednesday, October 16, 2024 8:49 PM
To: Alex Frawley <AFrawley@susmangodfrey.com>; KVP-OAI <KVPOAI@keker.com>; openaicopyrightlitigation.lwteam@lw.com; OpenAICopyright <OpenAICopyright@mofo.com>
Cc: #NewYorkTimes-Microsoft-FDBR <NewYorkTimes-Microsoft-FDBR@faegredrinker.com>; NewYorkTimes_Microsoft_OHS@orrick.com; NYT-AI-SG-Service@simplelists.susmangodfrey.com; dailynews-ai-rfem@rothwellfigg.com
Subject: RE: NYT / OAI's Responses to RFPs

EXTERNAL Email

Hi Alex,

Thanks for those responses. We’re glad to resolve the dispute over RFP 46. OpenAI is available to meet and confer regarding RFPs 18, 26, 27, 44, 45, 47, 48, 51, 52, 74-81, and 84-85 on Friday between 12pm-2pm PT / 3pm-5pm ET. Please let us know what time works best for you.

Looking forward to discussing the issues you raised in your emails.

Best,
Max

MAX LEVY

Associate | Morrison & Foerster LLP
755 Page Mill Road | Palo Alto, CA 94304-1018
P: +1 (650) 813-4063

From: Alex Frawley <AFrawley@susmangodfrey.com>

Sent: Wednesday, October 16, 2024 3:41 PM

To: KVP-OAI <KVPOAI@keker.com>; openaicopyrightlitigation.lwteam@lw.com; OpenAICopyright <OpenAICopyright@mofo.com>; Levy, Max I. <MLevy@mofo.com>

Cc: #NewYorkTimes-Microsoft-FDBR <NewYorkTimes-Microsoft-FDBR@faegredrinker.com>; NewYorkTimes_Microsoft_OHS@orrick.com; NYT-AI-SG-Service@simplelists.susmangodfrey.com; dailynews-ai-rfem@rothwellfigg.com

Subject: RE: NYT / OAI's Responses to RFPs

External Email

Hi Max,

Responses below in red.

Best,
Alex

From: Levy, Max I. <MLevy@mofo.com>

Sent: Wednesday, October 16, 2024 6:50 AM

To: Alex Frawley <AFrawley@susmangodfrey.com>; openaicopyrightlitigation.lwteam@lw.com; OpenAICopyright <OpenAICopyright@mofo.com>; KVP-OAI <KVPOAI@keker.com>

Cc: #NewYorkTimes-Microsoft-FDBR <NewYorkTimes-Microsoft-FDBR@faegredrinker.com>; NewYorkTimes_Microsoft_OHS@orrick.com; NYT-AI-SG-Service@simplelists.susmangodfrey.com

Subject: RE: NYT / OAI's Responses to RFPs

EXTERNAL Email

Hi Alex, following up on Monday's call, please see our responses to your 10/9 email below. In tandem with yesterday's email, we note that the responses below address all of the RFPs at issue in The Times's proposed agenda for the upcoming status conference (RFPs 18, 26-27, 44, 45-48, 53, 74-81 and 84-85). We believe our response moots these issues. If you disagree, please let us know. **We disagree, as explained below.**

RFP 18

As we noted in our responses and objections, the request for "documents concerning the removal of any content . . . from Training Datasets" (emphasis added) is overly broad, unduly burdensome, and disproportionate to the needs of the case. As written, this RFP could, for example, sweep in the deletion of an extra character in a given word to ensure it's spelled correctly. In light of your explanation that you are seeking documents relevant to your DMCA claim, in response to this Request, OpenAI agrees to conduct a reasonable search for and produce any non-privileged documents in its possession, custody, or control concerning whether to remove from datasets on which the relevant models were trained the following categories of information with respect to copyrighted works: authorship, date of publication, publishing entity, title, copyright notices, and terms and conditions for use.

We do not agree to the "whether to remove" language. But does OpenAI agree to this revised version of the request: "documents concerning the removal of copyright management information before, during, or after training of Defendants' Generative AI

Model(s), including: information regarding authorship, date of publication, publishing entity, title, copyright notices, and terms and conditions for use of works”? This revision accounts for the concern you raised about deletion of a word.

RFP 44

We continue to believe that our prior proposal is sufficient. But in the interest of compromise, OpenAI will agree to conduct a reasonable search for non-privileged documents in its possession, custody, or control discussing OpenAI's approach to paywalls with respect to the relevant models and products.

This offer is insufficient, and the News Plaintiffs stand by the position we took for this RFP in the draft position statement for the agenda.

RFPs 45-48, 51-52

OpenAI's investigation remains ongoing as we continue to determine what responsive, nonprivileged information exists regarding each of these RFPs. Nonetheless, based on the results of that investigation to date, in an effort to resolve the disputes, OpenAI responds as follows to each of the RFPs. As the investigation continues, we will supplement the responses as needed.

- **RFP 45:** OpenAI is reviewing this request and will follow up in due course. When can we expect a response? We served these RFPs on May 22, so the investigation should have been completed a long time ago.
- **RFP 46:** OpenAI agrees to conduct a reasonable search for non-privileged documents in its possession, custody, or control describing how the relevant products generate suggested follow-up queries for users. This offer resolves the dispute for this RFP, except for the global dispute about the “relevant” products.
- **RFP 47:** OpenAI agrees to conduct a reasonable search for non-privileged documents in its possession, custody, or control analyzing or summarizing OpenAI user click-through data for the relevant products. This offer is insufficient. The RFP is not limited to click-through the data. What about the other user engagement metrics and analytics addressed in the RFP, including duration, the number of prompts or queries that a user inputs per session, fallback rates, and bounce rates?
- **RFP 48:** OpenAI agrees to conduct a reasonable search for non-privileged documents showing the number of daily active users of each of the relevant products, from 2021 to the present, in the manner and at the frequency such information is kept in the ordinary course of business. Will this data include a break down for new users and repeat users on each day, from 2021 to present, and will it include dates of subscription and changes in subscriber numbers over time? Also, any resolution on this RFP will be subject to the global dispute about the “relevant” products.
- **RFP 51:** OpenAI agrees to conduct a reasonable search for non-privileged documents in its possession, custody, or control analyzing or summarizing OpenAI user click-through data for the relevant products. OpenAI further agrees to conduct a reasonable search for non-privileged documents in our possession, custody, or control discussing OpenAI's efforts, if any, to modify clickthrough rates for the relevant products. We do not agree to limit the RFP to “analyzing or summarizing.” OpenAI should produce documents “concerning” the issues addressed in the RFP. As for those issues, please confirm that “clickthrough rates,” as used in your response, will encompass information about users navigating away from Defendants’ products and services to other websites. Also, any resolution on this RFP will be subject to the global dispute about the “relevant” products.
- **RFP 52:** OpenAI is reviewing this request and will follow up in due course. When can we expect a response? We served these RFPs on May 22, so the investigation should have been completed a long time ago.

RFPs 74-81:

[REDACTED]

RFP 84-85: Contrary to Plaintiffs' assertion, OpenAI has not "refused to produce documents in response to these Requests." While OpenAI did refuse to produce documents in response to RFP No. 85 (because "it seeks "a]ll documents' regarding an unbounded universe of third-party websites with no connection to this litigation"), OpenAI specifically agreed to meet and confer regarding the scope of RFP No. 84. We appreciate The Times's efforts to articulate relevance, but are still unclear how requests concerning DNYuz and "similar" unnamed websites are proportional to The Times's needs of the case. We remain eager to meet and confer to better understand The Times's position. **We stand by these requests, including for the reasons set forth in the draft agenda position statements. When are you available Thursday or Friday to meet and confer? As part of these discussions, how many documents hit on the term "*DNYuz*" for OpenAI's custodians?**

MAX LEVY

Associate | Morrison & Foerster LLP

755 Page Mill Road | Palo Alto, CA 94304-1018

P: +1 (650) 813-4063

From: Alex Frawley <AFrawley@susmangodfrey.com>

Sent: Wednesday, October 9, 2024 10:40 AM

To: Levy, Max I. <MLevy@mofo.com>; openaicopyrightlitigation.lwteam@lw.com; OpenAICopyright <OpenAICopyright@mofo.com>; KVP-OAI <KVPOAI@keker.com>

Cc: #NewYorkTimes-Microsoft-FDBR <NewYorkTimes-Microsoft-FDBR@faegredrinker.com>; NewYorkTimes_Microsoft_OHS@orrick.com; NYT-AI-SG-Service@simplelists.susmangodfrey.com

Subject: RE: NYT / OAI's Responses to RFPs

External Email

Hi Max,

Can we meet and confer tomorrow or Friday about the RFPs in my October 4 email? We would like to resolve what we can resolve prior to filing our agenda for the conference next week.

Relatedly, we wanted to follow up on additional outstanding issues regarding OpenAI's responses to The Times's Second and Third Sets of RFPs. We hope to resolve these without court intervention as well and would like to meet and confer tomorrow or Friday.

1. RFP 18

- This RFP seeks documents concerning the removal of any content from Training Datasets, including information regarding authorship, date of publication, publishing entity, title, copyright notices, and terms and conditions for use of works. These documents are relevant to the DMCA claims. OpenAI has not yet agreed to produce any documents in response to this request, only offering to meet and confer.

2. RFPs 44, 53

- These RFPs seek documents relating to how Defendants' products enable users to circumvent paywalls. OpenAI has agreed to produce documents "sufficient to show ChatGPT's approach to paywalls," but that is insufficient. Among other issues, these Requests are relevant to The Times's contributory infringement claim, as well as Defendants' arguments, raised in their

motions to dismiss, that “normal people do not use OpenAI’s products” to elicit copyrighted content. Dkt. 52 at 2. These requests are also relevant to The Times’s harm allegations. These requests are also relevant because they refute OpenAI’s argument that OpenAI is permitted to use Times content insofar as it is publicly available. OpenAI should produce non-custodial and custodial documents in response to these requests.

3. RFPs 45-48, 51-52

- These RFPs seek documents relating to user engagement metrics for Defendants’ products, including logs that analyze user queries and activities. OpenAI has not yet agreed to produce any documents in response to these requests, only offering to meet and confer. We understand from our discussions about Interrogatory No. 8 that OpenAI has investigated what information it has pertaining to “user engagement, output, user feedback, and analytics for the relevant products.” We also understand from our discussions about Interrogatory No. 9 that OpenAI is investigating the existence of information relating to “click-through rates or analyses of user queries and outputs for ChatGPT.” We’d like to meet and confer to discuss what documents OpenAI has found as part of these investigations and whether there is a path to resolving these RFPs.

4. RFPs 74-81

- [REDACTED]

5. RFPs 84-85

- These RFPs seek documents concerning OpenAI’s knowledge of websites like DNYuz that copy publisher content, and OpenAI’s use of DNYuz and similar websites. OpenAI has so far refused to produce documents in response to these Requests, which are relevant to, among other issues, The Times’s claim of willful infringement. OpenAI conceded the relevance of third-party websites that copy publisher content by stating that regurgitation is “more common when particular content appears more than once in training data, like if pieces of it appear on lots of different public websites.”

From: Levy, Max I. <MLevy@mofo.com>

Sent: Tuesday, October 8, 2024 9:47 AM

To: Alex Frawley <AFrawley@susmangodfrey.com>; KVP-OAI <KVPOAI@keker.com>;
openaicopyrightlitigation.lwteam@lw.com; OpenAICopyright <OpenAICopyright@mofo.com>

Cc: #NewYorkTimes-Microsoft-FDBR <NewYorkTimes-Microsoft-FDBR@faegredrinker.com>;
NewYorkTimes_Microsoft_OHS@orrick.com; NYT-AI-SG-Service@simplelists.susmangodfrey.com

Subject: RE: NYT / OAI's Responses to RFPs

EXTERNAL Email

Hi Alex,

We’ve received The Times’s latest proposals regarding RFPs 17, 20, 21, 22, 23, 26, 27, 28, 32, 53, 54, 55, 61, 65, 66, and 69 following our August 26th and 29th meet and confers. We are reviewing your positions and investigating your questions, including exploring potential ways to narrow or resolve the outstanding disputes. We will follow up to your

email in writing within a week. From there, we can figure out the most efficient way to proceed on any remaining disputes.

Best,
Max

MAX LEVY

Associate | Morrison & Foerster LLP

755 Page Mill Road | Palo Alto, CA 94304-1018

P: +1 (650) 813-4063

From: Alex Frawley <AFrawley@susmangodfrey.com>

Sent: Friday, October 4, 2024 3:42 PM

To: KVP-OAI <KVPOAI@keker.com>; openaicopyrightlitigation.lwteam@lw.com; OpenAICopyright <OpenAICopyright@mofo.com>

Cc: #NewYorkTimes-Microsoft-FDBR <NewYorkTimes-Microsoft-FDBR@faegredrinker.com>; NewYorkTimes_Microsoft_OHS@orrick.com; NYT-AI-SG-Service@simplelists.susmangodfrey.com

Subject: NYT / OAI's Responses to RFPs

External Email

Counsel,

I write to follow up on open issues related to OpenAI's responses to The Times's Second Set of RFPs, including as discussed on the parties' August 26 and 29 meet-and-confer calls. We would like to continue to meet and confer about this set of RFPs. Please let us know when you are available on Tuesday or Wednesday. Given the upcoming conference, we ask that OpenAI provide its final positions on the below issues during the meet and confer.

Category II.A from The Times's July 31 letter, addressing RFPs for which OpenAI initially refused to produce any documents

RFP 53

In its August 20 letter, OpenAI said that it "already agreed to produce responsive, non-privileged documents in its possession, custody, or control, if any, sufficient to show ChatGPT's approach to paywalls that it locates pursuant to a reasonable and diligent search (RFP No. 44)" and that, "for the avoidance of doubt, OpenAI will confirm that as part of that production, it will conduct a reasonable search for and produce responsive, non-privileged documents, if any, related to the Browse feature's interactions with paywalls and the role, if any, that feature played in the decision to suspend Browse with Bing and/or Bing Chat, to the extent such documents exist."

As discussed on the call, The Times believes that OpenAI should also produce custodial documents in response to this Request, particularly given OpenAI's public acknowledgement that this decision involved concerns about the feature reproducing publisher content. Will OpenAI produce custodial documents?

RFP 54

In Its August 20 letter, OpenAI said that it will produce non-privileged documents related to its statement that “The Times paid someone to hack OpenAI’s products.”

As discussed on the call, this response is insufficient. OpenAI should also produce documents relating to the other statements identified in the RFP, including (1) OpenAI’s January 8, 2024 publication titled “OpenAI and Journalism,” <https://openai.com/index/openai-and-journalism/>, and (2) various statements made by Sam Altman at the World Economic Forum in Switzerland in January 2024, including his statement that “we actually don’t need to train on [Times] data,” his statement that OpenAI wanted to pay The Times “a lot of money to display their content,” his statement that “the [regurgitation] problem is not as easy as it sounds in a vacuum. I think we can get that number down and down and down, quite low, ” and his statement that “I think there’s going to be great new ways to consume and monetize news and other published content,” as memorialized in this article: <https://www.cnn.com/2024/01/18/openai-ceo-on-nyt-lawsuit-ai-models-dont-need-publishers-data.html>. Each of these statements – and any others like them – relate to critical issues in this litigation.

RFP 55

OpenAI in its August 20 letter asked whether The Times will agree to limit the scope of this RFP to the “specific employees and board members” listed in the RFP. We discussed that Request during the parties’ August 26 and 29 calls. The Times can clarify that it will agree to limit the RFP to the employees and board members listed in the RFP, plus these additional employees: Peter Deng, John Schulman, and Greg Brockman (who we understand is on leave). We served these RFPs before these employees left and/or took leave. For the avoidance of doubt, OpenAI should also produce documents relating to subsection VI of the RFP: Microsoft’s involvement in the hiring and firing of employees of any OpenAI entity.

RFP 61

On the parties’ August 26 and 29 calls, OpenAI stated that it is unsure whether any joint defense agreement with Microsoft is privileged. Please clarify OpenAI’s position. OpenAI also asked whether The Times would consider limiting this Request to agreements with Microsoft in effect before The Times filed this lawsuit. To the extent that any post-filing agreement does not relate to Microsoft or OpenAI’s obligations as to this lawsuit, The Times is willing to so limit the RFP. Will OpenAI now agree to produce documents in response to this RFP, as narrowed?

RFP 65

During the parties’ August 26 and 29 calls, OpenAI claimed this Request is overbroad in part because OpenAI is a GenAI company. The Times is willing to narrow this Request to documents concerning OpenAI’s public relations strategy as it pertains to (1) its models’ and products’ reliance on or use of publisher and copyright-protected content, and (ii) issues related to copyright law, including those related to any element of OpenAI’s anticipated fair use defense.

RFP 66

OpenAI during our August 26 and 29 calls reiterated its view that it does not believe the requested documents are in any way relevant to the case. The Times continues to disagree. One specific issue we discussed was whether OpenAI has templates of emails or other communications that it has sent or will send to potential investors or actual investors. Can you tell us whether such templates exist? That information would enable us to evaluate your assertion from your August 20 letter that the RFP is “unduly burdensome” because it would require OpenAI to produce “almost any email to any potential investor.” Aside from producing these templates, The Times will narrow this Request to documents relating to (i) financial data, including projections; (ii) OpenAI’s plans and strategies to commercialize and/or monetize its products;

(iii) OpenAI's non-profit structure; (iv) OpenAI's use of copyrighted content; and (iv) potential or existing lawsuits or investigations that are related to any of the issues in the Complaint.

RFP 69

OpenAI on the August 26 and 29 calls continued to assert that it does not understand the relevance of this Request. The Times continues to disagree. We can also clarify that this Request seeks documents about the 2022 change in OpenAI's practices regarding public access to OpenAI's governing documents, financial statements, and conflict of interest rules, as reflected in this article (which was cited in the RFPs): <https://www.wired.com/story/openai-scrapped-promise-disclose-key-documents/>. The Times can also clarify that for now, it is willing to narrow this Request to documents "sufficient to show" why OpenAI made this change. If OpenAI agrees to this narrowed Request, The Times will review those productions before determining whether any additional document productions on this topic are warranted.

Category II.B from The Times's July 31 letter, addressing RFPs for which OpenAI's responses are incomplete or require clarification

RFP 17

OpenAI confirmed on the August 26 and 29 calls that it has not yet produced all "agreements relating to text training data used to train the models for ChatGPT," including because OpenAI is going through the process of providing notice to third parties about these productions. Please let us know which specific agreements are forthcoming and confirm that all such agreements will be produced by October 11, subject to any third-party moving for a protective order.

RFPs 20, 21, 22

For RFP 20, OpenAI agreed to produce documents "within its possession, custody, and control, if any, sufficient to show the total number of records and tokens in OpenAI's training datasets that can be located pursuant to a reasonable and diligent search." As discussed on the parties' calls, The Times reiterates its objection to OpenAI's use of "if any" to qualify this response to the RFP. At a minimum, after conducting a reasonable investigation, OpenAI should clarify whether it found and produced responsive documents, and if not, OpenAI should supplement its response to confirm that it lacks any responsive documents.

Based on the parties' calls, including on August 26 and 29, we also understand OpenAI's position to be that it does not have any documents responsive to the portion of this Request seeking documents sufficient to show the total number of records and tokens in the Training Datasets that contain (i) Times Content, (ii) Journalism content, and (iii) copyrighted content, and that OpenAI's position is that the parties must discover this information through the separately negotiated training data inspection process. The Times understands OpenAI to take the same position with respect to RFPs 21 and 22, which seek, respectively, documents sufficient to show each source of records in Defendants' Training Datasets, and documents sufficient to identify each article of Times Content contained in Defendants' Training Datasets. Is it OpenAI's position that it has conducted a reasonable search and does not have any documents responsive to these Requests?

Finally, OpenAI's August 22 letter asserts that OpenAI also trained the models based on "other information generated by OpenAI for the purpose of AI training." At the next meet and confer, we'd like you to elaborate more on that statement, including whether "information generated by OpenAI" was produced by a model containing Times content.

RFP 23

On the August 26 and 29 calls, we clarified what we meant by the terms “contemplated” and “steps.” Specifically, this Request is meant to encompass both actual training and operating activities involving Times content as well as documents concerning other options that OpenAI explored for training with Times or Journalism content aside from the methods and processes OpenAI ultimately adopted. We understand that OpenAI is considering whether to produce documents in response to this Request.

RFPs 26-27

These Requests seek documents concerning OpenAI’s knowledge of and use of web crawlers and related tools to access Times content, such as but not limited to GPTBot, ChatGPT-User, OAI-SearchBot, CommonCrawl bot, as well as any crawlers used by Microsoft or developers of Custom GPTs. The Times seeks documents concerning the operation of such crawlers (including technical documentation, source code, and logs) and OpenAI’s use of or reliance on such crawlers to access The Times’s sites to train its Generative AI Models or power its Generative AI Products and Services.

RFP 28

On the August 26 and 29 calls, we clarified that this Request seeks documents relevant to The Times’s DMCA claim. For example, The Times seeks documents relating to whether OpenAI intended to conceal the source of its training data, including but not limited to any intent to conceal its reliance upon Times Content. The Times relatedly seeks documents about the technical process by which OpenAI prepared materials for use as training data, including documents concerning any alterations made to the training data, such as the removal or alteration of CMI. This Request applies both to OpenAI’s development of the Generative AI Models as well as the Generative AI Products and Services.

We understand that OpenAI is considering whether to produce documents in response to this Request, as clarified.

RFP 32

As stated on the August 26 and 29 calls, The Times reiterates its position that OpenAI’s “sufficient to show” limitation is improper. The Times is also confused as to why OpenAI is objecting to producing custodial documents when OpenAI has already run search terms related to the training process, including [book* w/5 (source* OR train* OR data* OR collection)] and, more recently, OpenAI included terms like “cycle” and “epoch” in its search term proposal. OpenAI should produce custodial documents in response to this RFP. To aid in the process, The Times can further clarify, as it did on the calls, that some of the topics in which it is interested include any training “recipes” as well as the sequence in which content is loaded and used during training.

Alexander P. Frawley | Susman Godfrey LLP

One Manhattan West, 50th Floor | New York, NY 10001

212.729.2044 (office) | 917.599.6613 (cell)

afrawley@susmangodfrey.com | www.susmangodfrey.com

=====

This message may be confidential and privileged. Use or disclosure by anyone other than an intended addressee is prohibited. If you received this message in error, please delete it and advise the sender by reply email. Learn about Morrison & Foerster LLP's [Privacy Policy](#).

This message may be confidential and privileged. Use or disclosure by anyone other than an intended addressee is prohibited. If you received this message in error, please delete it and advise the sender by reply email. Learn about Morrison & Foerster LLP's [Privacy Policy](#).

.

This message may be confidential and privileged. Use or disclosure by anyone other than an intended addressee is prohibited. If you received this message in error, please delete it and advise the sender by reply email. Learn about Morrison & Foerster LLP's [Privacy Policy](#).

.

To unsubscribe from this list please go to <https://https://simplelists.susmangodfrey.com/subs/>

This message may be confidential and privileged. Use or disclosure by anyone other than an intended addressee is prohibited. If you received this message in error, please delete it and advise the sender by reply email. Learn about Morrison & Foerster LLP's [Privacy Policy](#).

.

To unsubscribe from this list please go to
<https://simplelists.susmangodfrey.com/confirm/?u=flGOkiG8u2jiaPQ0BbCXTK2YLSBMieHZ>